Using GIS and spatial modeling to examine active travel potential in a university town

Greg Rybarczyk, Ph.D.
Dept. of Earth and Resource Science
University of Michigan-Flint

OVERVIEW

- Introduction – Universities and Active Transportation
- Collecting the right data
- Exploratory Analysis + Spatial modeling
- Implications for stakeholders

INTRODUCTION

- **Background**
  - Universities are major trip generator-attractors
    - Traffic congestion
    - Personal Safety
    - Pollution
    - Attractiveness
    - Faculty retention
  - Universities are microcosms of society

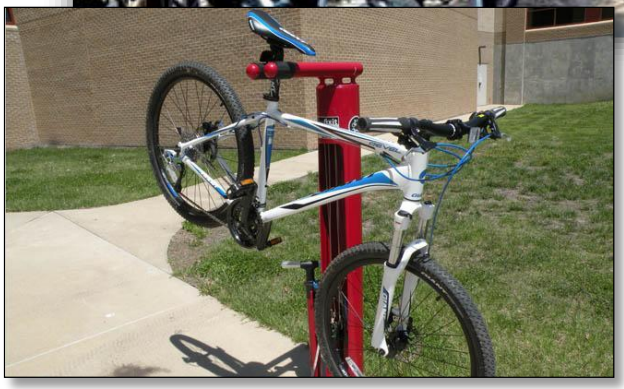
INTRODUCTION

- **Universities and Active Transportation**
- Benefits of increasing walking, bicycling, and mass transit utilization among a university population
  - Reduce environmental externalities
  - Student recruitment
  - Beacon of sustainability
  - Educate next generation of planners/decision makers

INTRODUCTION

- **Universities and Active Transportation**
- University utilize Travel Demand Analysis (TDM)
  - Demand vs. Supply
  - Transit subsidies
  - Bicycle facilities
  - Traffic Calming
  - Programming
  - City-wide partnerships
  - Reducing parking demand
  - Courses

INTRODUCTION



KU CYCLING

INTRODUCTION

- What about spatial and aspatial factors at the trip-origin?
  - Neighborhood conditions
  - Family constraints
  - Socioeconomic conditions
  - Limited access to school
  - Traffic conditions
  - Distance
  - Weather
  - Social stigmas
  - Etc…

INTRODUCTION

- Research goals:
  - Identify the factors that may cause a mode shift to active transportation at the trip start
- Required tools:
  - Survey instrument
  - GIS
  - Spatially explicit model

DATA

Survey Instrument

- Gender, university classification, home address
  - Distance from UM-Flint
  - Mode Share to & on UM-Flint campus
    - Auto
    - Bicycle
    - Bus
    - Walking
    - Scooter, etc.
- Conditions affecting mode choice
  - 13 Interventions to *increase biking*
  - 7 Barriers to *decrease biking*
  - 8 Interventions to *increase walking*

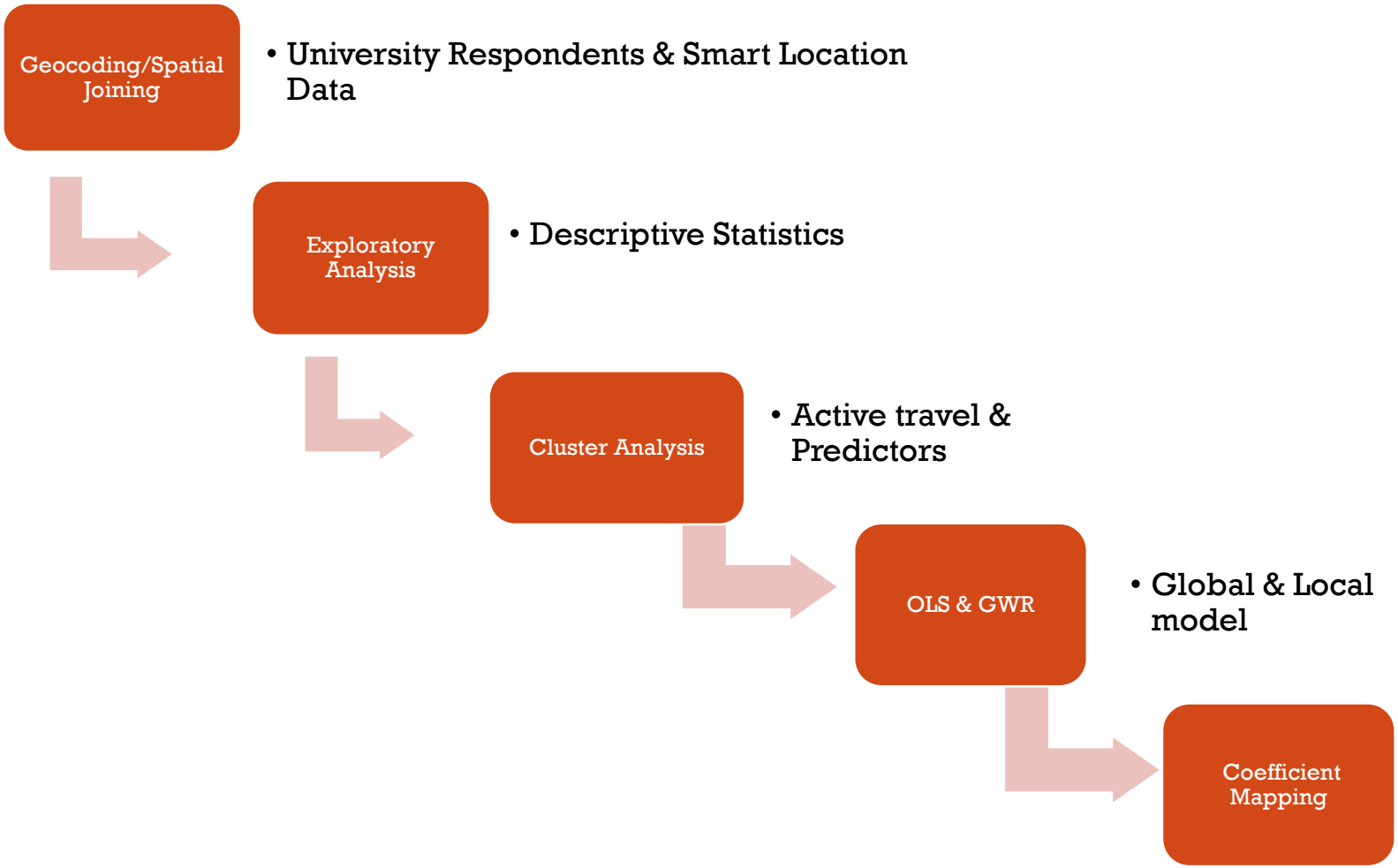
DATA

Neighborhood Context

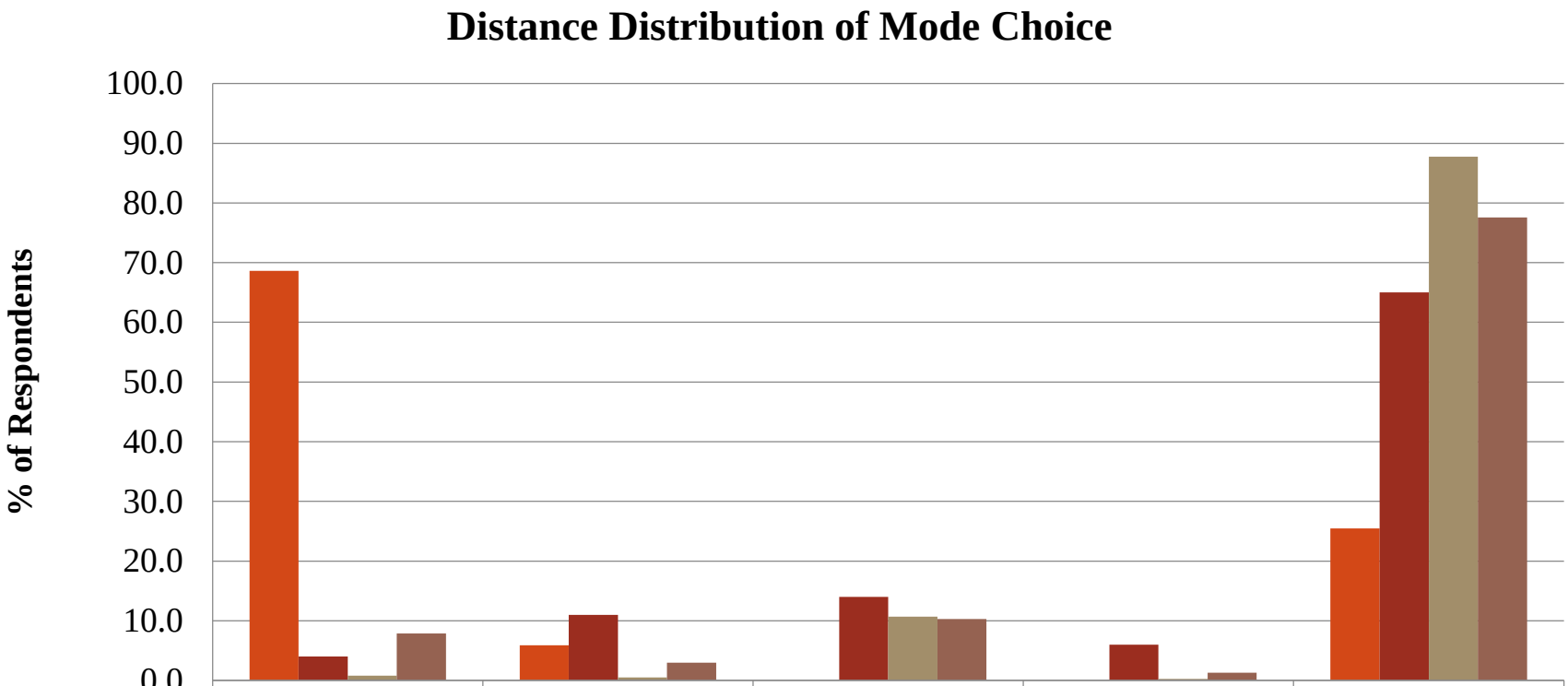
- 2010 Environmental Protection Agencies "Smart Location Database"
  - Census Block Group
    - Employment
    - Land-use Diversity
    - Demographics
    - Density
    - Design
  - Transportation Environment
    - Bicycle and Pedestrian Crashes

Density	Gross residential density (housing units per acre) on unprotected land
Diversity of land use	Employment and housing entropy
Urban design	Street intersections per square mile High-speed road network density
Transit service	Aggregate transit service frequency, afternoon peak period Distance to nearest transit stop
Destination accessibility by transit*	Working-age population within a 45-minute transit commute
Destination accessibility by car	Jobs within a 45-minute drive
Demographics	Percentage of households with no car, 1 car, or 2 or more cars Percentage of workers that are low, medium, or high wage (by home and work locations)
Employment	Employment totals broken down by 5-tier classification scheme

MODELING APPROACH



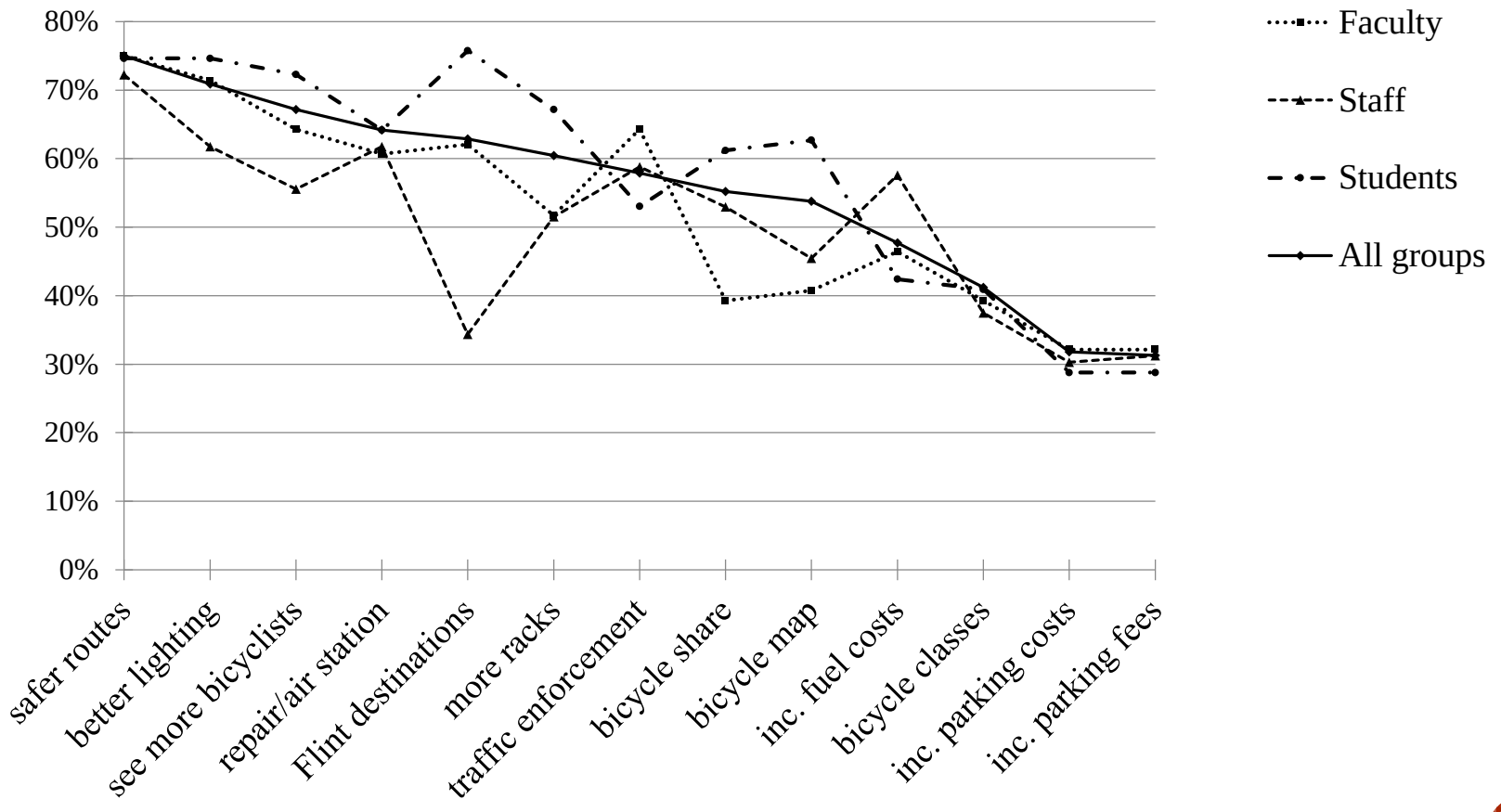
EXPLORATORY ANALYSIS



	Walk %	Bike %	Car Pool%	Bus %	SOV %
<1 mile	68.6	5.9	0.0	0.0	25.5
1-5 miles	4.0	11.0	14.0	6.0	65.0
> 5 miles	0.8	0.5	10.7	0.3	87.8
OVERALL	7.9	3.0	10.3	1.3	77.6

EXPLORATORY ANALYSIS

Bicycling facilitators for those who are > 5 miles from campus

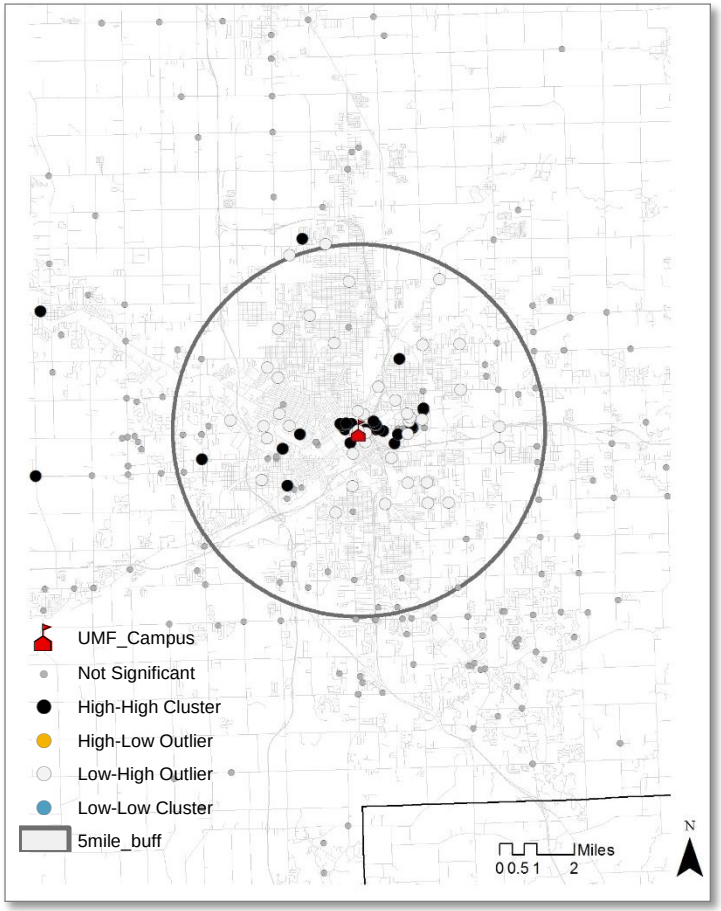


EXPLORATORY ANALYSIS

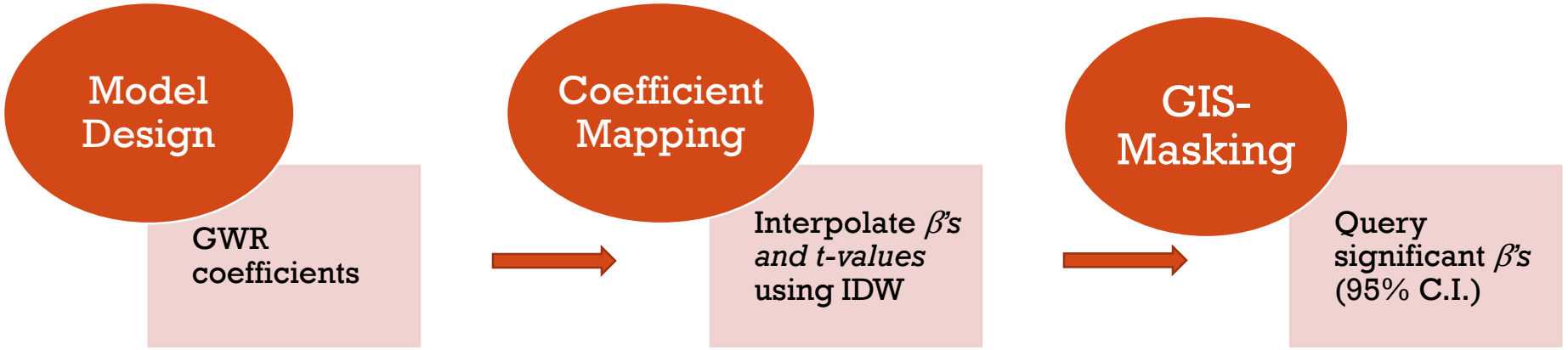
Variable	Moran's I	Pattern	p-value
BikeWalkBus	1.355	Clustered	0.00
DistUMFlint	0.498	Clustered	0.00
Male[a]	-0.038	Random	.513
Faculty[b]	0.124	Clustered	.022
Student[b]	0.099	Clustered	.068
% Households with Workers	1.301	Clustered	0.00
% Zero Car Households	1.300	Clustered	0.00
Res. Density	1.268	Clustered	0.00
Auto-orientated Facility Density	2.120	Clustered	0.00
Regional Job/Pop Diversity	1.492	Clustered	0.00
Bike/Ped Crash Density	0.839	Clustered	0.00

[a] = control - female
[b] = control - staff

Local Moran's I – Active Travel Mode

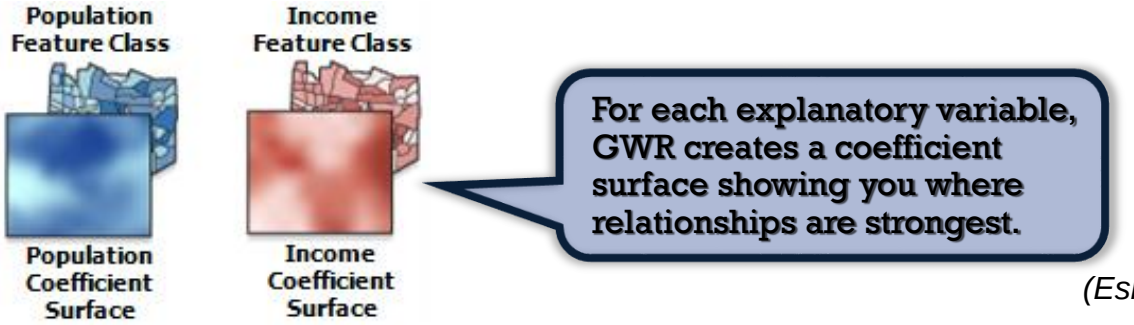


SPATIAL MODELING



SPATIAL MODELING

- Two different types of modeling approaches that can be implemented in GIS
- OLS (Ordinary Least Squares)
  - Global regression model
  - One equation, calibrated using data from all features
  - Relationships are fixed
  - Does not account for spatial heterogeneity *(Wen et al., 2010)*
- GWR
  - Local regression model
  - One equation for every feature, calibrated using data from nearby features



(Esri, 2010)

SPATIAL MODELING

$$\textbf{GWR} = \quad y_i = \beta_{io} + \sum_{k=1}^{p} \beta_{ik} x_{ik} + \epsilon_i$$

Where:

y_i = dependent variable at location i ($i = 1, 2, \ldots..,n$, where n is the number of observations),

x_{ik} = independent variable of the kth parameter at location i,

β_{ik} = estimated kth parameter at location i for the GWR model,

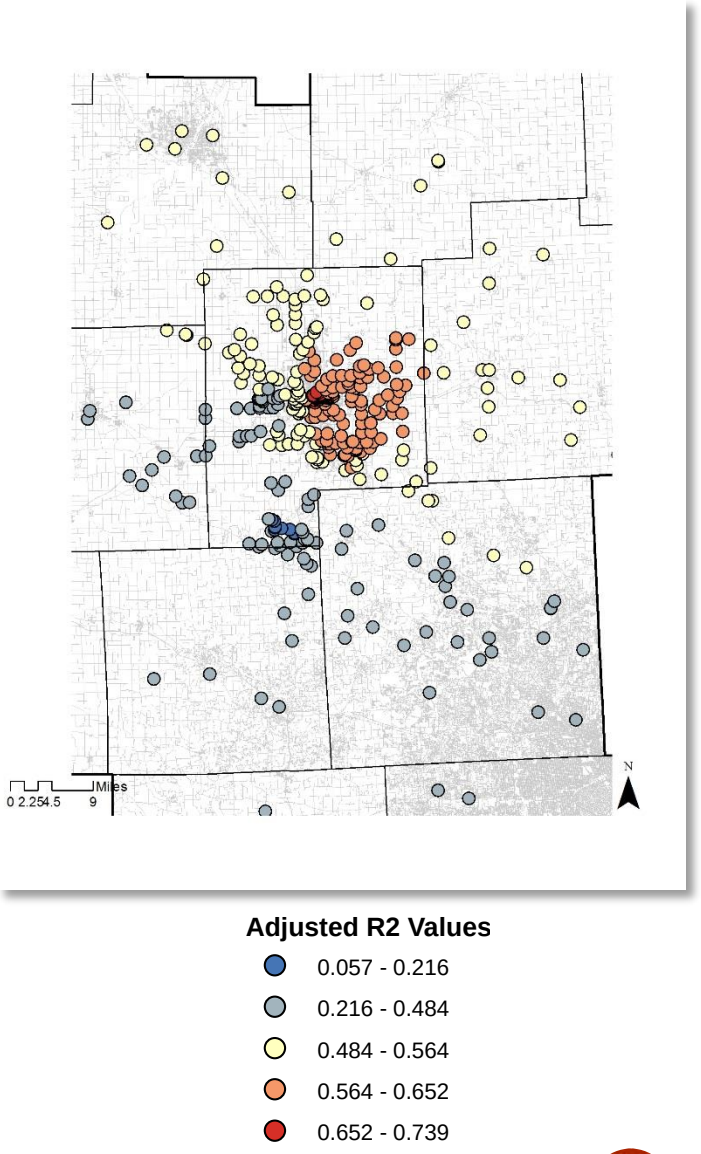
β_k = estimated kth parameter for the OLS model,

ϵ_i = error term at location i , and

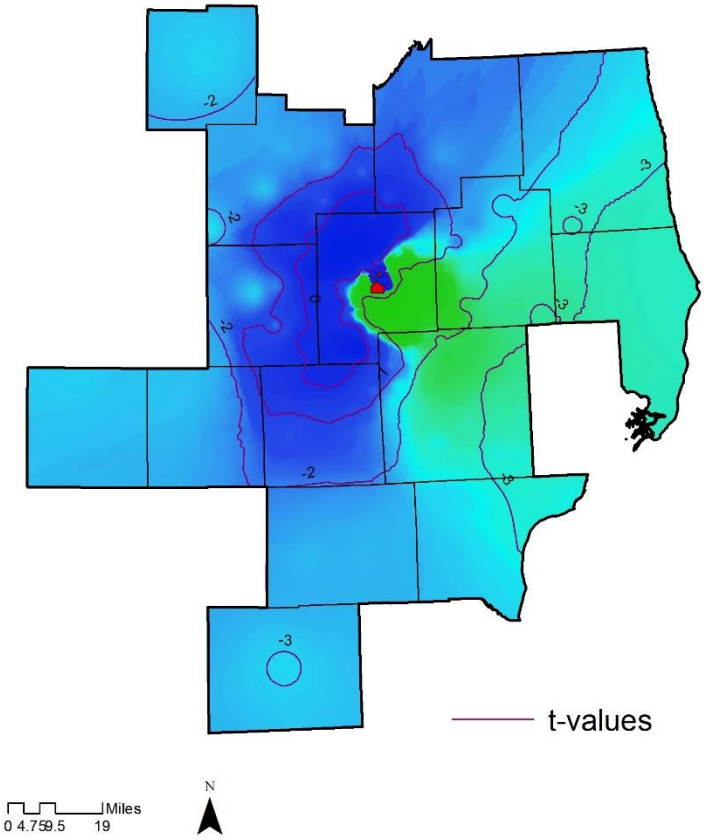
p = number of parameters

SPATIAL MODELING

- Typical GWR outputs include:
- Best-fit (R^2)
- Parameter Estimates (magnitude of influence, -+)
- T-values (distribution of significance, -+)
  - Mapping only the parameter estimate can be misleading (where is it significant?)
  - T-values have been displayed as contour lines over the parameter estimates
  - Can be "messy" or differences may be too large to interpret



SPATIAL MODELING

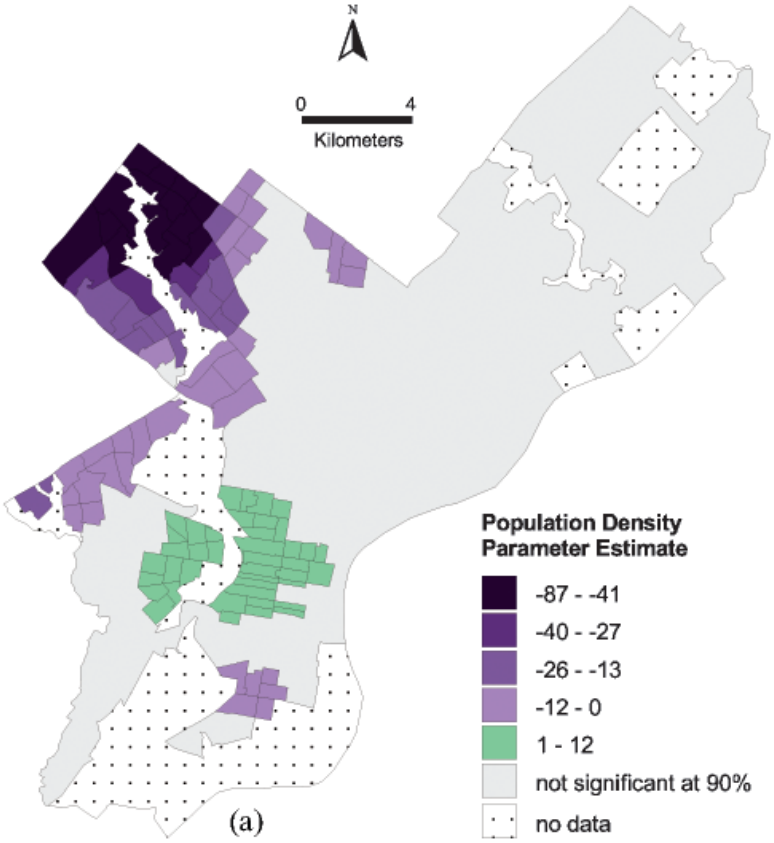


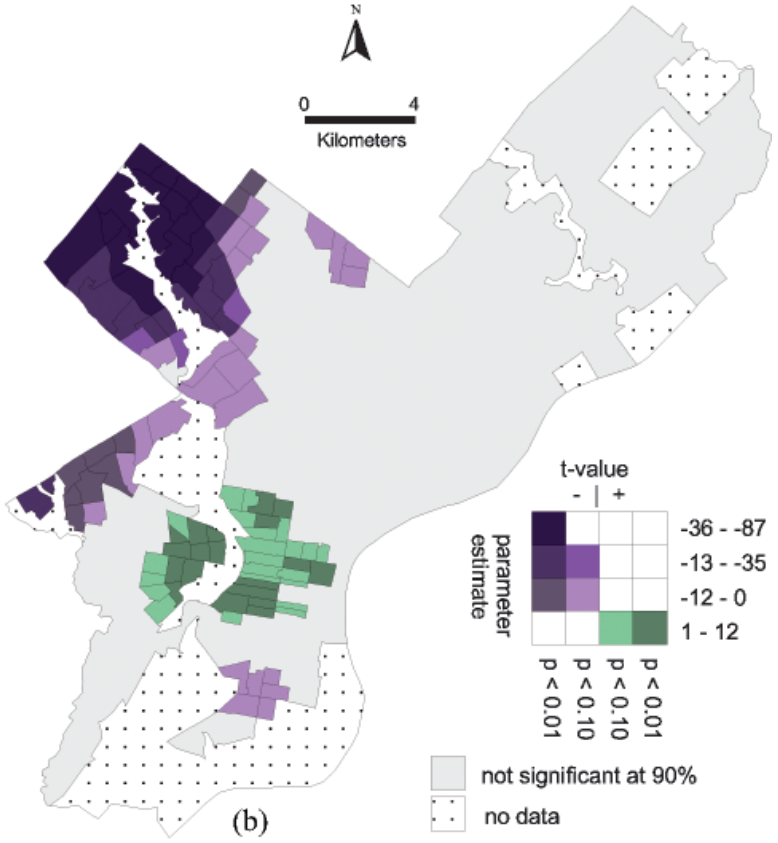
- T-values overlaid onto parameter estimates (standardized coefficients)
- Provides a vague understanding of where estimates are statistically significant

(Matthews, 2012)

SPATIAL MODELING

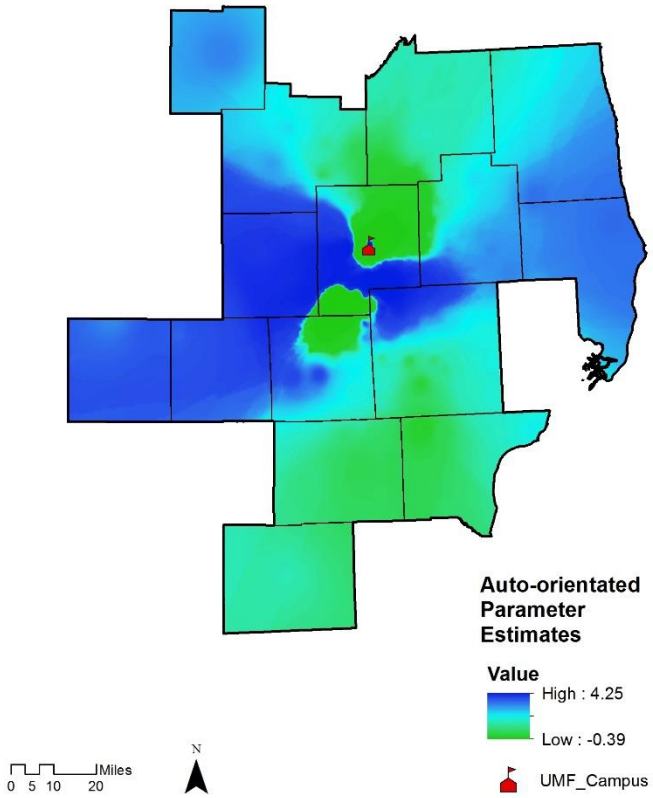
Creative symbologies can be used to link estimates to t-values

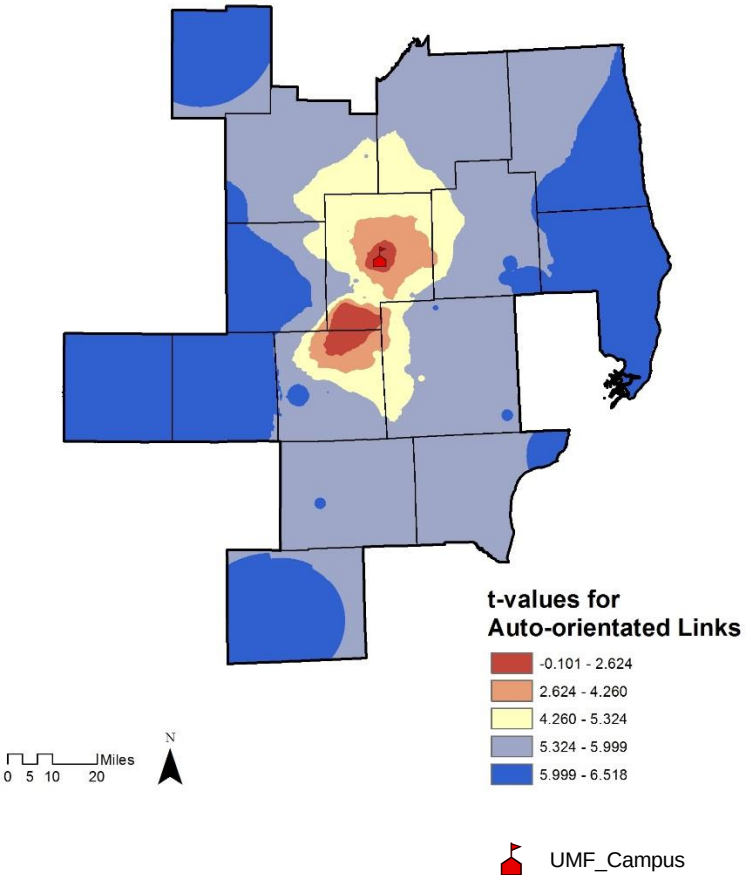




(Mennis, 2006)

SPATIAL MODELING





Where are the parameter estimates *significantly* affecting active mode choice potential?

SPATIAL MODELING

- **Integrating parameter estimates and statistical significance**

1. Create surface of local estimates (β's) for each explanatory variable using Inverse Distance Weighting (IDW)
2. Create surface of t-values for each explanatory variable Inverse Distance Weighting
   - T-values become the z's
3. Classify t-value surface for statistical significance
   - Choose 3 classes and manually change ranges (break values) to -1.96 and 1.96
4. Choose "no-color" for the 1st and 3rd classes (transparent - unique). Choose white for the 2nd class (opaque)
   - This is used as a mask in ArcGIS
5. Produce bivariate color scheme to display parameter estimates

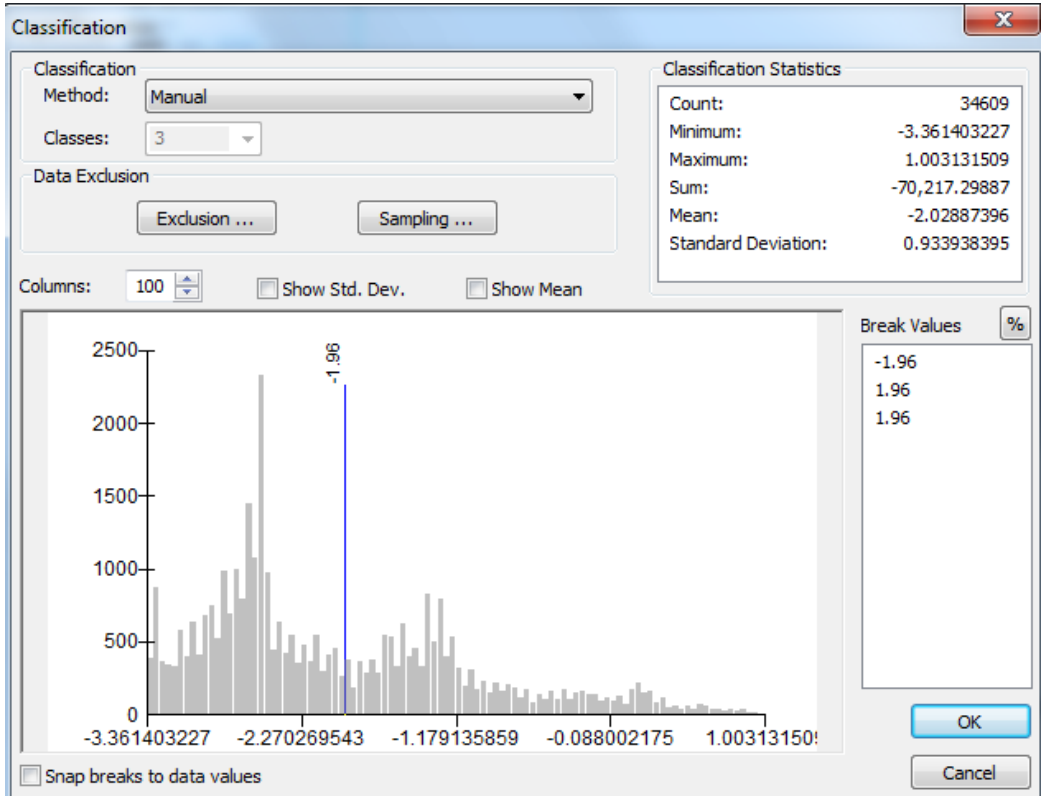
(Matthews et al., 2012)

SPATIAL MODELING

T-value partitions

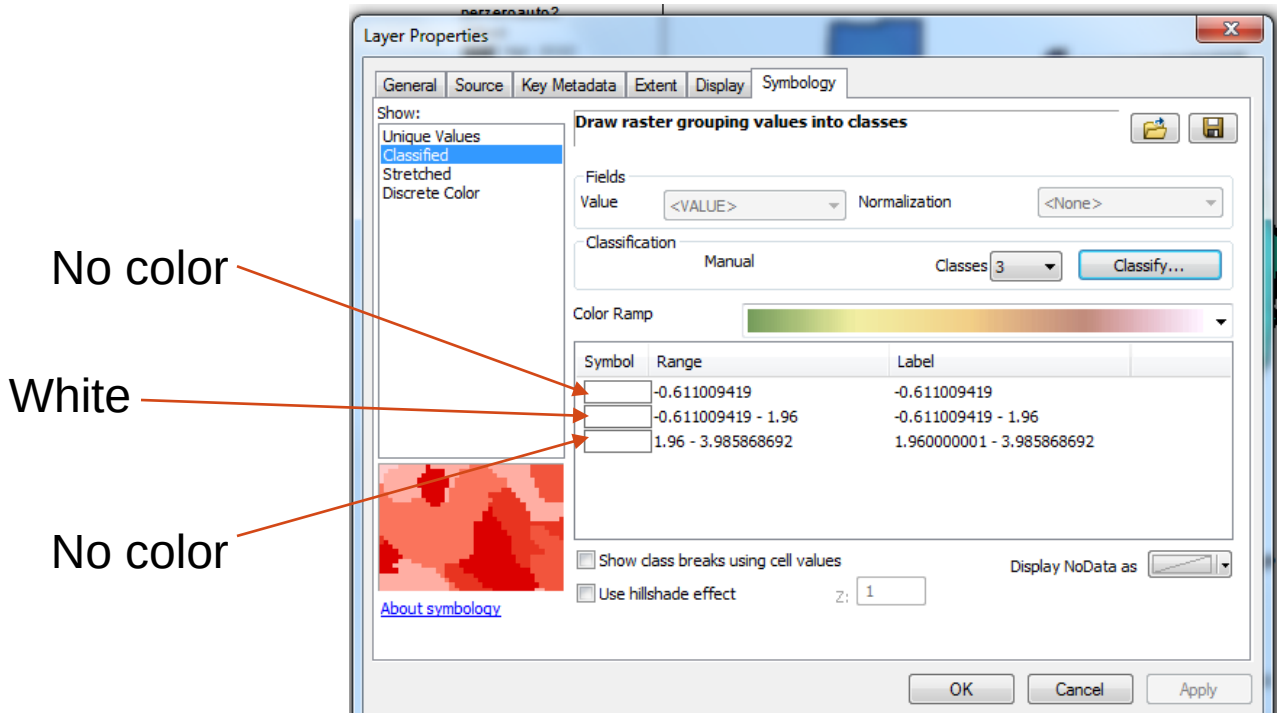
1.96 is the approximate value of the 95% point of the normal distribution used in statistics.

Other critical values can be queried:
99%, 2.58
90%, 1.645

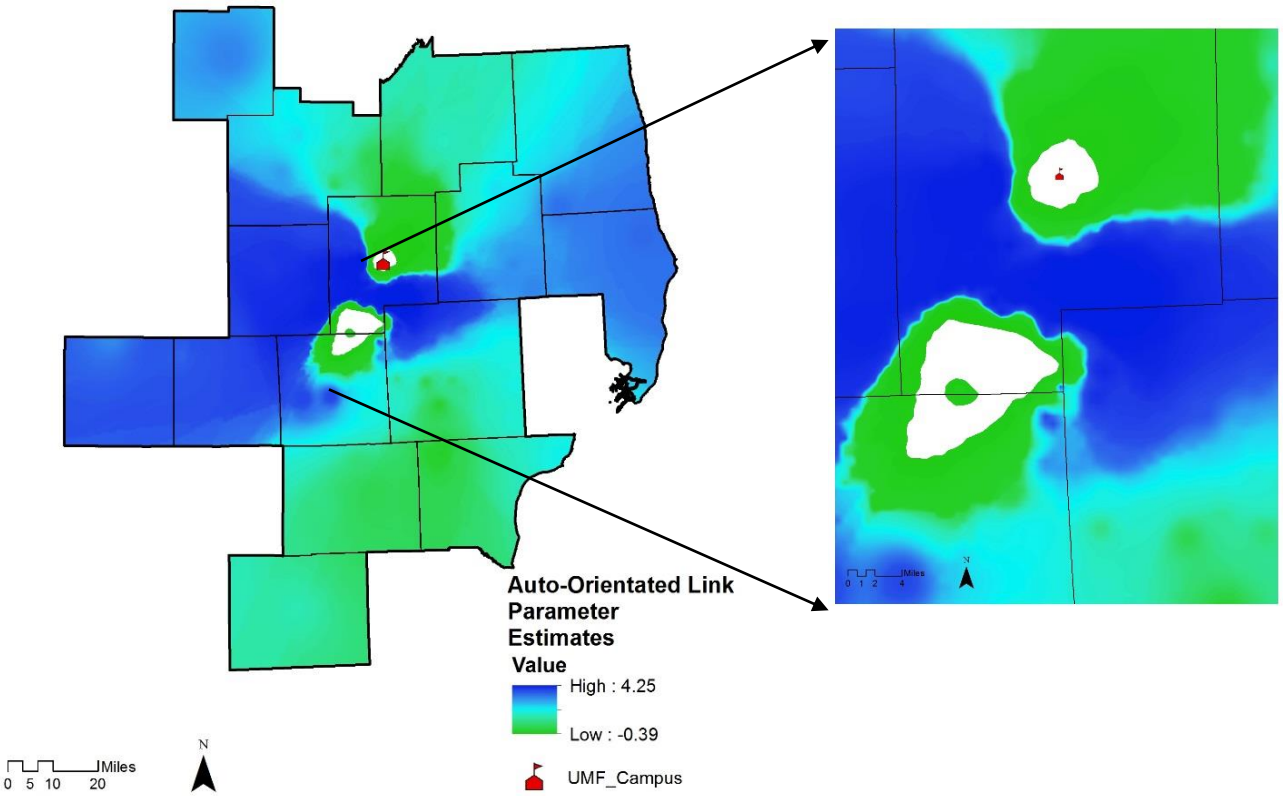


SPATIAL MODELING

- Setting up the significance mask

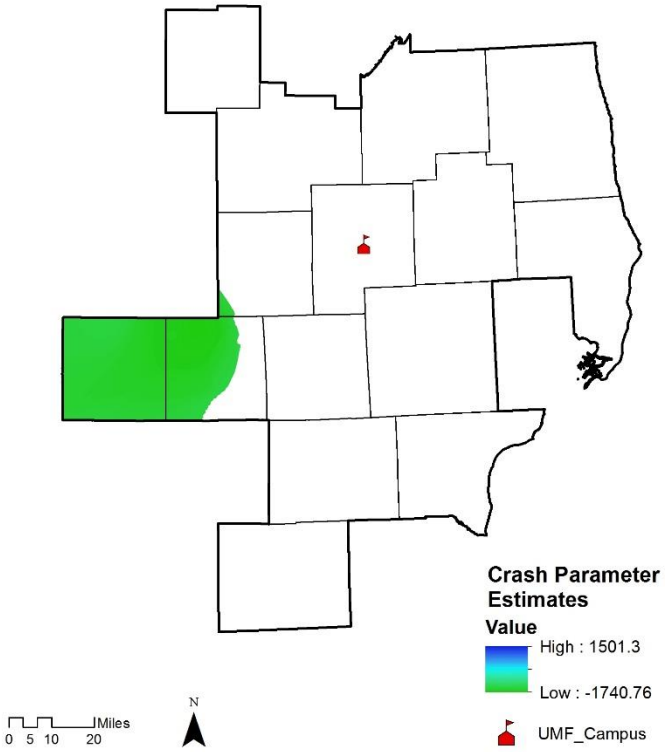


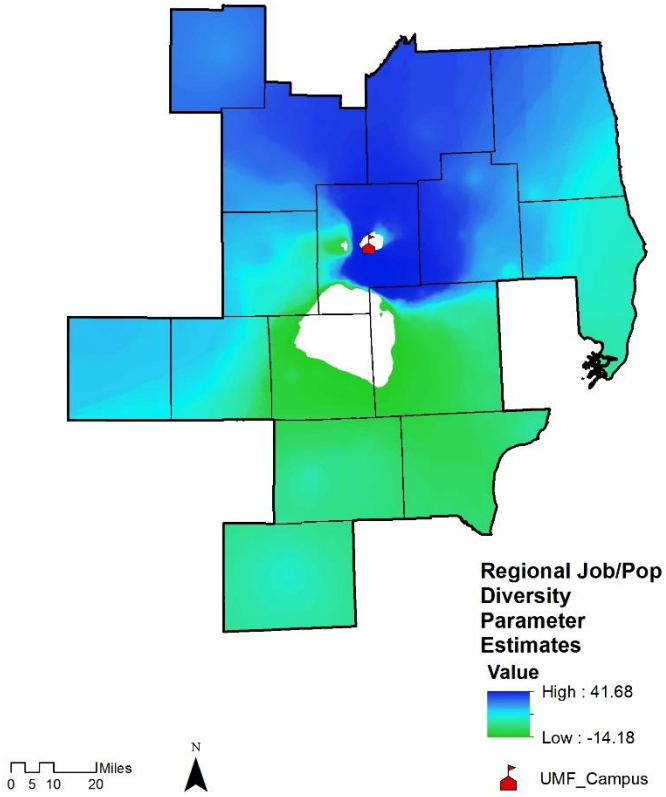
SPATIAL MODELING



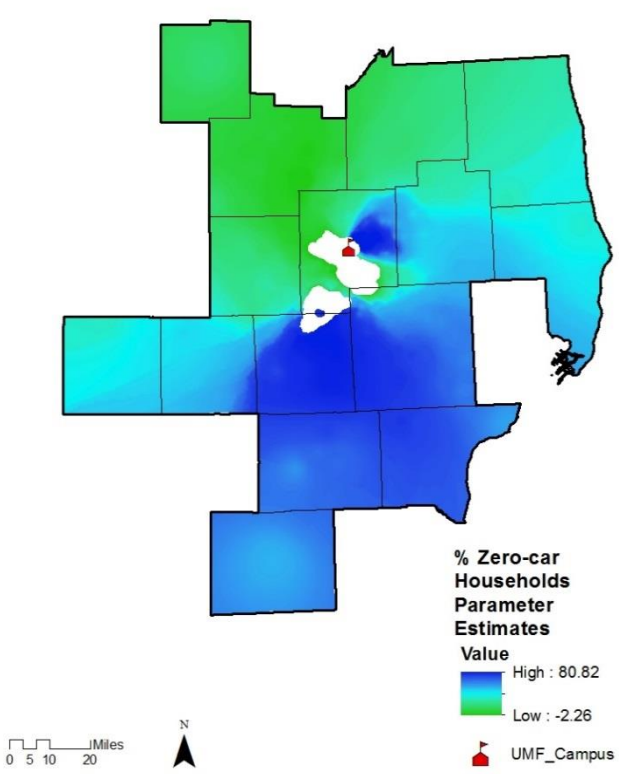
- We can now **visualize** where the parameter estimates may be affecting *local* active transportation mode choices, positively or negatively

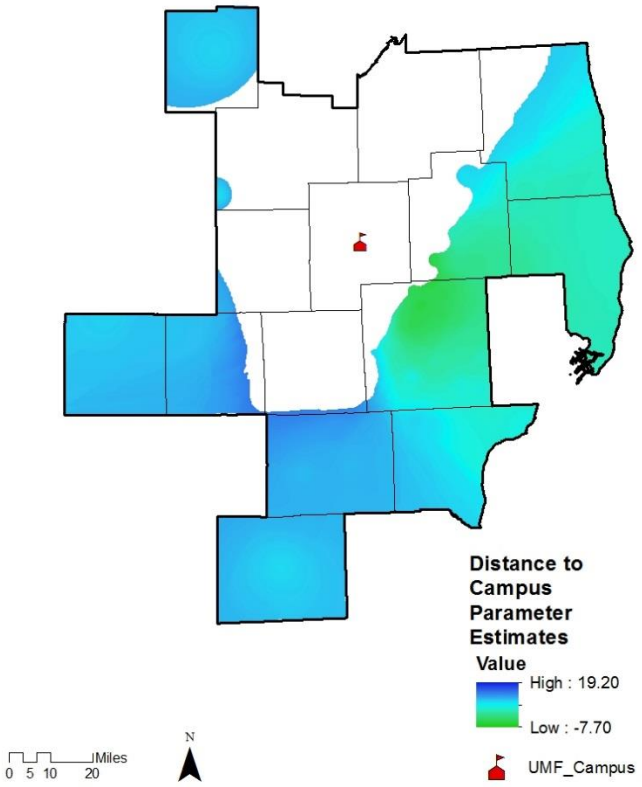
SPATIAL MODELING





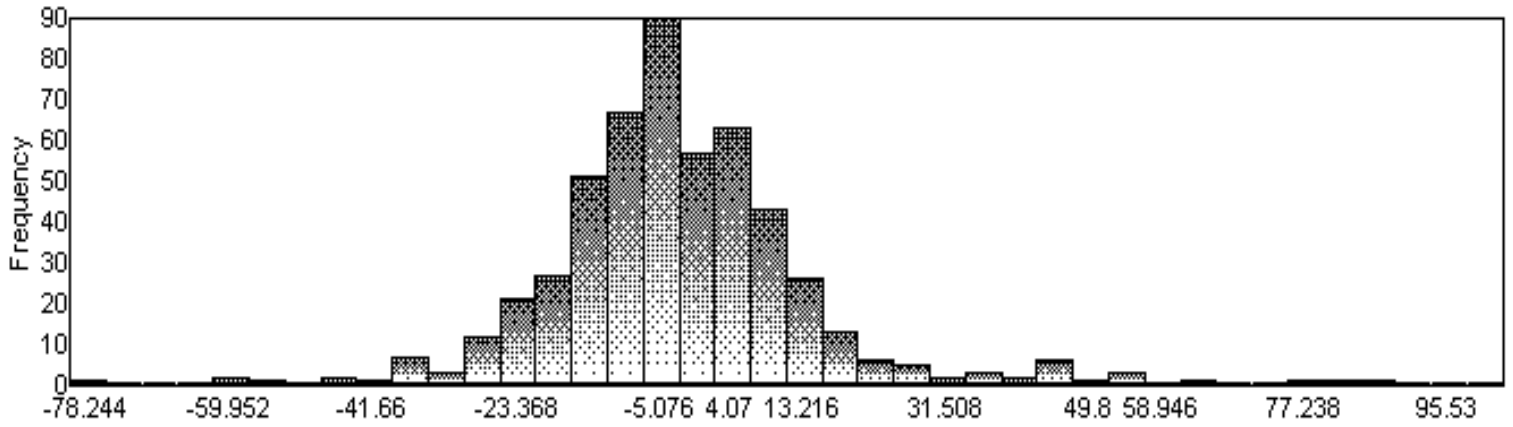
SPATIAL MODELING





SPATIAL MODELING

- **GWR Model Residuals**



SPATIAL MODELING

- **Overall Model Comparisons:**

Overall Model Diagnostics		
Estimated diagnostics of OLS and GWR models (n=520)		
Variables	**OLS**	**GWR**
R^2	.473	.625
Adjusted R^2	.464	.587
Akaike's Information Criterion (AICc)	4671.26	4595.38
F statistic	45.71	16.35
Sigma	11.00	48.83
Residual sum of squares	231381.1	169547.67
p-value	<.001	<.001

IMPLICATIONS

- Job/Housing diversity
  - Positively influences north = increased mixed-use campus development?
- Travel safety (crashes)
  - Significant negative affect on active travel west of campus = educational programming efforts needed?
- Household car density
  - # cars per household increases has positive affect south of campus?
- Distance
  - Negative influences east of campus = closer park and ride locations?
  - No significant effect north
- GWR outperformed OLS model

IMPLICATIONS

- Universities are in the transportation business!

- Examining what will propel faculty, staff, and students to use mass transit, walking, and bicycling can have far ranging affects on the university and host

- Spatially explicit modeling approaches and novel symbolization techniques can highlight where "best-practices" are needed

Thank You